IN THE CLAIMS:

Claims 1-67. (canceled)

Claim 68.   (new)   A computer-implemented method for simultaneous visualization of disparate data types, the method comprising:

(1)   selecting a set of attributes associated with an object, the attributes selected comprising a text data type and one other data type chosen from

a biopolymer sequence data type,

a numerical data type, and

a categorical data type;

(2)   creating a high dimensional vector representing the object by applying transformation operations to the selected attributes; and

(3)   projecting the high dimensional vector thereby visualizing the object based on the attributed selected;

wherein the transformation operations for the attributes of the text data type comprise:

(a)   semantically filtering a set of documents in a database to extract a set of semantic concepts, to improve an efficiency of a predictive relationship to its content, based on at least one of word frequency, overlap and topicality;

(b)   defining a topic set, said topic set being characterized as the set of semantic concepts which best discriminate the content of the

2

documents containing them, said topic set being defined based on at least one of word frequency, overlap and topicality;

(c)   forming a matrix with the semantic concepts contained within the topic set defining one dimension of said matrix and the semantic concepts contained within the filtered set of documents comprising another dimension of said matrix;

(d)   calculating matrix entries as the conditional probability that a document in the database will contain each semantic concept in the topic set given that it contains each semantic concept in the filtered set of documents; and

(e)   providing said matrix entries from step (d) for creating the high dimensional vector.

Claim 69.   (new)   A computer-implemented method for simultaneous visualization of disparate data types, the method comprising:

(1)   selecting a set of attributes associated with an object, the attributes selected comprising a biopolymer sequence data type and one other data type chosen from
        a text data type,
        a numerical data type, and
        a categorical data type;

(2)   creating a high dimensional vector representing the object by applying transformation operations to the selected attributes; and

(3)   projecting the high dimensional vector thereby visualizing the object based on the attributed selected;

3

wherein the transformation operations for the attributes of the biopolymer sequence data type comprise:

    (i)   comparing a sequence of each biopolymer material to a sequence of each other biopolymer material to provide respective comparison results;

    (ii)   arranging the comparison results in a square matrix indexed by the plurality of biopolymer materials; and

    (iii)   providing the square matrix entries for creating the high dimensional vector.

Claim 70.   (new)   The computer-implemented method of claims 69, wherein the attributes selected in step (1) comprise a text data type and a biopolymer sequence data type.

Claim 71.   (new)   The computer-implemented method of claims 69, wherein the transformation operations for the attributes of the text data type comprise:

    (a)   semantically filtering a set of documents in a database to extract a set of semantic concepts, to improve an efficiency of a predictive relationship to its content, based on at least one of word frequency, overlap and topicality;

    (b)   defining a topic set, said topic set being characterized as the set of semantic concepts which best discriminate the content of the documents containing them, said topic set being defined based on at least one of word frequency, overlap and topicality;

4

(c) forming a matrix with the semantic concepts contained within the topic set defining one dimension of said matrix and the semantic concepts contained within the filtered set of documents comprising another dimension of said matrix;

(d) calculating matrix entries as the conditional probability that a document in the database will contain each semantic concept in the topic set given that it contains each semantic concept in the filtered set of documents; and

(e) providing said matrix entries from step (d) for creating the high dimensional vector.

Claim 72. (new) The computer-implemented method of claim 70, wherein the attributes selected in step (1) comprise a text data type, a biopolymer sequence data type, and one other data type chosen from a numerical data type and a categorical data type.

Claim 73. (new) The computer-implemented method of claim 70, wherein the attributes selected in step (1) comprise a text data type, a biopolymer sequence data type, a numerical data type, and a categorical data type.

Claim 74. (new) A computer-implemented method for simultaneous visualization of disparate data types, the method comprising:

(1) selecting a set of attributes associated with an object, the attributes selected comprising any three data types chosen from
a text data type,
a biopolymer sequence data type,

5

a numerical data type, and

a categorical data type;

(2)   creating a high dimensional vector representing the object by applying transformation operations to the selected attributes; and

(3)   projecting the high dimensional vector thereby visualizing the object based on the attributed selected;

wherein the transformation operations for the attributes of the text data type, if selected, comprise:

(a)   semantically filtering a set of documents in a database to extract a set of semantic concepts, to improve an efficiency of a predictive relationship to its content, based on at least one of word frequency, overlap and topicality;

(b)   defining a topic set, said topic set being characterized as the set of semantic concepts which best discriminate the content of the documents containing them, said topic set being defined based on at least one of word frequency, overlap and topicality;

(c)   forming a matrix with the semantic concepts contained within the topic set defining one dimension of said matrix and the semantic concepts contained within the filtered set of documents comprising another dimension of said matrix;

(d)   calculating matrix entries as the conditional probability that a document in the database will contain each semantic concept in the topic set given that it contains each semantic concept in the filtered set of documents; and

6

(e) providing said matrix entries from step (d) for creating the high dimensional vector;

and wherein the transformation operations for the attributes of the biopolymer sequence data type, if selected, comprise:

(i) comparing a sequence of each biopolymer material to a sequence of each other biopolymer material to provide respective comparison results;

(ii) arranging the comparison results in a square matrix indexed by the plurality of biopolymer materials; and

(iii) providing the square matrix entries for creating the high dimensional vector.

Claim 75. (new) A computer-readable medium containing software for performing the method of any one of claims 67-74.

Claim 76. (new) A device adapted to perform the method of claim 75.

Claim 77. (new) The method of any of claims 67-74, wherein said application of transformation application on said selected attributes produces a vector representation of said object in correspondence with a uniform data structure.

Claim 78. (new) A computer-readable medium containing software for performing the method of claim 77.

Claim 79. (new) A device adapted to perform the method of claim 77.

7

Claim 80.　(new)　The method of claim 77, further comprising using said representation to identify cluster groups of related objects.

Claim 81.　(new)　A computer-readable medium containing software for performing the method of claim 80.

Claim 82.　(new)　A device adapted to perform the method of claim 80.